Claims

1. Method for automatic online detection and classification of anomalous objects in a data stream, especially comprising datasets and / or signals,

characterized in

a) the detection of at least one incoming data stream (1000) containing normal and anomalous objects,

b) the automatic construction (2100) of a geometric representation of normality (2200) of the incoming objects of the data stream (1000) at a time t_1 subject to at least one predefined optimality condition, especially the construction of a hypersurface enclosing a finite number of normal objects,

c) the online adaptation of the geometric representation of normality (2200) in respect to at least one received object at a time $t_2 > t_1$, the adaptation being subject to at least one predefined optimality condition,

d) the online determination of a normality/anomality classification (2300) for received objects at t_2 in respect to the geometric representation of normality (2200),

e) the automatic classification of normal objects and anomalous objects based on the generated normality classification (2300) and generating a data set describing the anomalous data for further processing, especially a visual representation.

2. Method according to claim 1, characterised in that the geometric representation of normality (2200) is a parametric boundary hypersurface using the enclosure of the

minimal volume or the minimal volume estimate among all possible surfaces as an optimality condition.

3. Method according to claim 2, characterised in that the hypersurface is constructed in the space of original measurements of least one incoming data stream (1000) or in a space obtained by a nonlinear transformation thereof.

4. Method according to at least one preceding claim, characterised in that the optimality condition, used to construct the parametric boundary hypersurface, is a pre-defined condition, especially the one based on an expected fraction η of anomalous objects, or a condition, dynamically adaptable to the data stream.

5. Method according to at least one preceding claim, characterised in that the anomalous objects are determined as the ones lying outside of the geometrical representation of normality (2200), especially the parametric boundary hypersurface enclosing the normal objects.

6. Method according to at least one preceding claim, characterized in that dynamic adaptation of the geometric representation of normality (2200) comprises an automatic adjustment of parameters x_i of the geometric representation of normality (2200) to incorporate at least one new object while maintaining the optimality of the geometric representation of normality (2200).

7. Method according to at least one preceding claim, characterized in that the dynamic adaptation of the geometric representation of normality (2200) comprises an automatic adjustment of parameters x_i of the geometric representation of normality (2200) to remove the least-relevant object while maintaining the optimality of the geometric representation of normality (2200).

8. Method according to at least one preceding claim, characterized in that the smallest volume geometric representation of normality (2200) is maintained from an instance t_i after which the construction of the geometric representation of normality (2200) is feasible subject to the optimality condition.

9. Method according to at least one preceding claim, characterized in that the geometric representation of normality (2200) is generated with a Support Vector Machine method, generating a parametric vector x to describe the representation.

10. Method according to at least one preceding claim, characterised in that the temporal change of the geometrical representation of normality (2200), especially the temporal change of a parameter vector x of the geometrical representation of normality (2200) is stored for the evaluation of temporal trend in the data stream (1000).

11. Method according to at least on one preceding claim, characterised in that the geometric representation of normality (2200) is a sphere or any part thereof.

12. Method according to at least one preceding claim, characterized in that incoming data stream (1000) comprises data ~~packets in communication networks or~~ representations thereof.

13. Method according to at least one preceding claim, characterized in that the data objects comprises entries originating from the logging in process in at least one computer or representations thereof.

14. Method according to claim 12 or 13, characterized in that the determination of normality of the received data packets distinguishes normal incoming data stream from

anomalous data, especially sniffing attacks and / or denial of service attacks, whereby the means for automatic determining the normal and anomalous data generates a warning message.

15. A method according to any preceding claim, characterized in that, the method for construction and update of the geometric representation of normality (2200) in which the coordinate system in which the representation is constructed is fixed to some point in the data space or in the feature space.

16. A method according to claim 15, in which the center of coordinate system coincides with the center of mass of the data space (in the original or in the feature space)

17. A method according to claim 15 or 16, in which the decision on normality or anomality of an object is decided upon its norm in the data-centered (or feature-space-centered) coordinate system, or by the radius of the hypersphere centered at the center of the origin in the said coordinate system and encompassing the given objects.

18. A method according to one of the claims 15 to 17 in which the update of the representation includes the update of the coordinate system.

19. A method according to one of the claims 15 to 18 in which the update of coordinate system includes the update of the center of coordinates.

20 A method according to one of the claims 15 to 19 in which importation of the new object includes as a part the update of the norms of all objects in the working set so as to bring them in the new coordinate system corresponding to the expanded working set ("norm expansion").

21. A method according to one of the claims 15 to 20, in which removal of the object includes as a part the update of the norms of all objects in the working set so as to bring them in the new coordinate system corresponding to the contracted working set ("norm contraction")

22. System for automatic online detection and classification of anomalous objects in a data stream, especially comprising datasets and / or signals,

characterized by

a) a detection means for least one incoming data stream (1000) containing normal and anomalous objects,

b) an automatic online anomaly detection engine comprising

- an automatic construction means (2100) of a geometric representation of normality (2200) for the incoming objects of the data stream (1000) at a time t_1 subject to at least one predefined optimality condition, especially for the construction of a hypersurface enclosing a finite number of normal objects, with an automatic online adaptation means for the geometric representation of normality (2200) in respect to received at least one received object at a time $t_2 > t_1$, the adaptation being subject to at least one predefined optimality condition, and

- an automatic online determination means of a normality classification (2300) for received objects at t_2 in respect to the geometric representation of normality (2200).

c) an automatic classifcation means (4000) of normal objects and anomalous objects based on the generated normality classification (2300) and generating a data set describing

the anomalous data for further processing, especially a visual representation.